



UNIVERSIDAD DE LA RIOJA

TRABAJO FIN DE ESTUDIOS

Título

Cálculo del coeficiente de difusión y la viscosidad de un líquido mediante simulación computacional

Autor/es

ANDREA PUENTE IZQUIERDO

Director/es

RODRIGO MARTÍNEZ RUIZ y JUDITH MILLÁN MONEO ,

Facultad

Facultad de Ciencia y Tecnología

Titulación

Grado en Química

Departamento

QUÍMICA

Curso académico

2017-18



Cálculo del coeficiente de difusión y la viscosidad de un líquido mediante simulación computacional, de ANDREA PUENTE IZQUIERDO
(publicada por la Universidad de La Rioja) se difunde bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported.
Permisos que vayan más allá de lo cubierto por esta licencia pueden solicitarse a los titulares del copyright.



UNIVERSIDAD DE LA RIOJA

Facultad de Ciencia y Tecnología

TRABAJO FIN DE GRADO

Grado en Química

CÁLCULO DEL COEFICIENTE DE DIFUSIÓN Y LA VISCOSIDAD DE UN LÍQUIDO MEDIANTE SIMULACIÓN COMPUTACIONAL

Realizado por:

Andrea Puente Izquierdo

Tutelado por:

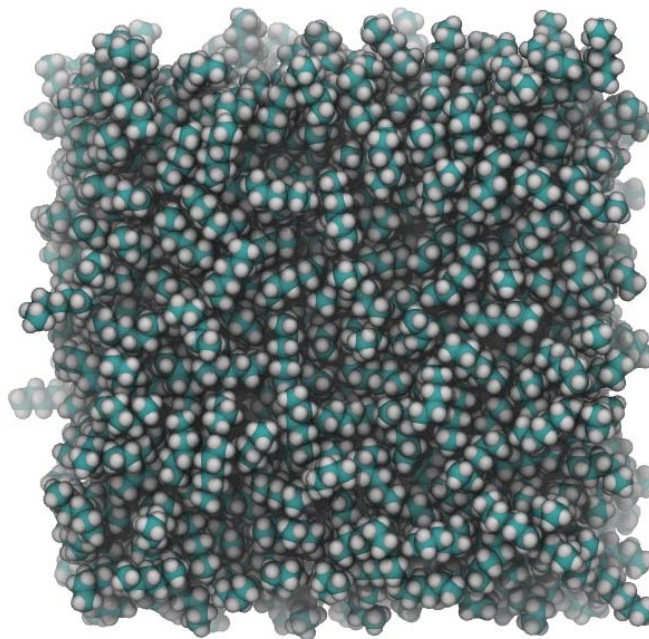
Rodrigo Martínez Ruiz

Judith Millán Moneo

Logroño, julio, 2018

TRABAJO FIN DE GRADO

CÁLCULO DEL COEFICIENTE DE DIFUSIÓN Y LA VISCOSIDAD DE UN LÍQUIDO MEDIANTE SIMULACIÓN COMPUTACIONAL



CALCULATION OF LIQUID'S DIFFUSION COEFFICIENT AND VISCOSITY THROUGH COMPUTATIONAL SIMULATION

Realizado por: Andrea Puente Izquierdo

Tutelado por: Rodrigo Martínez Ruiz

Judith Millán Moneo

ÍNDICE

RESUMEN	1
OBJETIVOS.....	3
INTRODUCCIÓN TEÓRICA	4
Mecánica Molecular	5
Dinámica Molecular	8
Minimización energética	8
Cálculo del coeficiente de difusión y de la viscosidad	13
PROCEDIMIENTO COMPUTACIONAL	17
RESULTADOS.....	22
CONCLUSIONES.....	31
ANEXOS	32
ANEXO 1: Archivos utilizados en el cálculo.....	33
ANEXO 2: Valores del coeficiente de difusión según el tamaño del sistema y la temperatura.....	42
ANEXO 3: Datos de ajuste para el coeficiente de difusión corregido	43

RESUMEN

Actualmente la Química Computacional permite desarrollar modelos computacionales que explican los complejos procesos quimicofísicos y ayudan a prever resultados. Esto es de vital importancia para minimizar los altos costes de los experimentos realizados en el laboratorio, ya que permite simular los análisis a realizar y observar cuáles son las condiciones óptimas.

En este Trabajo Fin de Grado se han estudiado ciertas propiedades de transporte del hexano aplicando la Dinámica Molecular. Con las simulaciones investigadas a las temperaturas 260, 280, 300, 320 y 340 K de sistemas compuestos por un número considerable moléculas de hexano se ha obtenido el coeficiente de autodifusión a cada una de las temperaturas de trabajo. A partir de estos valores de coeficientes de difusión, y aplicando la ecuación de Stokes-Einstein, se ha determinado la viscosidad del hexano a cada una de las temperaturas investigadas, comparando los resultados obtenidos con resultados experimentales.

Tras la realización de las simulaciones y el posterior análisis de los resultados obtenidos se ha podido comprobar la validez de este método de cálculo, aunque para futuras investigaciones sería conveniente la realización de mejoras, tales como optimizar la simulación de las moléculas de hexano consideradas como partículas esféricas así como ampliar el tamaño del sistema.

Los cálculos se realizaron con el *software* libre GROMACS. Parte de ellos se llevaron a cabo en los ordenadores de la Fase IV de la Facultad de Ciencia y Tecnología de la Universidad de La Rioja y, los cálculos computacionales más costosos, en el Laboratorio Computacional de Química Física de la misma universidad.

ABSTRACT

Nowadays Computational Chemistry helps to develop computational models that are able to explain complex physical chemist processes and foresee results. This is something vital to minimize costs of the experiments that take place in the laboratory and find out which conditions are the best for the process.

In this Bachelor Degree Thesis different transport properties of hexane have been studied applying Molecular Dynamics. With the investigated simulations at the temperatures 260, 280, 300, 320 and 340 K of systems made of a considerable amount of hexane molecules, the coefficient of self-diffusion was obtained at each of the working temperatures, comparing the obtained results with experimental ones. From the values for the diffusion coefficients and applying the Stokes-Einstein equation, it was calculated the viscosity of hexane for each investigated temperatures, comparing these values with the experimental data.

After carrying out the simulations and the subsequent analysis of the results, it has been possible to verify the validity of this method of simulation, although for future investigations it would be convenient to make improvements, such as spherical particles or expanding the size of the system.

The calculations were made with the GROMACS software. Some of them were carried out on phase IV computers from the Faculty of Science and Technology from University of La Rioja and, the most lengthy calculations were carried out in the Computational Laboratory of Physical Chemistry of the same university,

OBJETIVOS

- Utilizar las herramientas computacionales de la Mecánica y Dinámica Molecular para estudiar propiedades de transporte de un líquido profundizando en los conceptos adquiridos en las asignaturas Química Física III y Química Computacional Aplicada.
- Aprender a utilizar el sistema operativo Linux Ubuntu a través de la consola de comandos y el significado de los diferentes comandos introducidos.
- Aprender a construir sistemas y realizar simulaciones utilizando el programa de Dinámica Molecular GROMACS
- Aprender a extraer la información relevante de los distintos ficheros de salida del programa.
- Validar el campo de fuerzas utilizado y los parámetros de cálculo a través de la comparación con los resultados experimentales.
- Calcular el coeficiente de autodifusión del hexano con los resultados obtenidos de las simulaciones a diferentes temperaturas.
- Calcular la viscosidad del hexano líquido a través del coeficiente de autodifusión y la ecuación de Stokes-Einstein y comprobar la validez del método de cálculo aplicado.
- Comparar las propiedades de transporte calculadas con resultados bibliográficos.

INTRODUCCIÓN TEÓRICA

Para centrar este Trabajo Fin de Grado (TFG) tenemos que comprender el fundamento de la Química Computacional. Los químicos computacionales se dedican a los aspectos químicos del problema mediante el manejo y la optimización de herramientas relacionadas con la computación como el desarrollo de algoritmos mejorados para resolver ecuaciones particularmente difíciles en el desarrollo de la codificación, la visualización de los datos etc., ya sea como entrada o salida de un modelo¹.

En este TFG se van a aplicar dos herramientas esenciales en Química Computacional, la Mecánica Molecular y la Dinámica Molecular.

Estas dos metodologías van a ser empleadas para el estudio de dos propiedades de transporte fundamentales de un líquido. Las propiedades de transporte de un sistema (líquido o gas) hacen referencia a la transferencia de materia, energía y cantidad de movimiento entre este sistema y sus alrededores o una parte y otra del mismo sistema².

El coeficiente de difusión de una especie hace referencia al transporte de materia en el medio en el que se encuentra. Las moléculas de un líquido (o gas) puro se mueven y hacen que la concentración sea homogénea en el medio, el coeficiente de difusión tiene relación directa con la velocidad a la que se mueven estas.

Por otra parte la viscosidad se relaciona con la resistencia de un fluido a fluir, tiene relación con el transporte del momento lineal expresado en la diferencia de velocidades con las que las diferentes capas que forman un fluido se mueven.

La Dinámica Molecular describe el movimiento de las partículas, así, con la descripción adecuada de la energía potencial mediante la Mecánica Molecular, se han calculado los valores del coeficiente de autodifusión y de la viscosidad del hexano a diferentes temperaturas.

¹ C. J. Cramer, *Essentials of Computational Chemistry*. WILEY (West Sussex, 2002).

² I. N. Levine, *Fisicoquímica*, Mc-Graw Hill (Madrid, 2006).

Mecánica Molecular

La Mecánica Molecular hace uso de la física clásica para el estudio de la Superficie de Energía de Potencial (SEP) de un sistema. La energía potencial que nos proporciona la Mecánica Molecular viene definida mediante el denominado campo de fuerzas. Los campos de fuerzas son bases de datos en los que se encuentran los parámetros para describir las interacciones enlazantes y no enlazantes del sistema estudiado.³

En este TFG, en el que se trabajará con hexano (Figura 1), las interacciones enlazantes son:

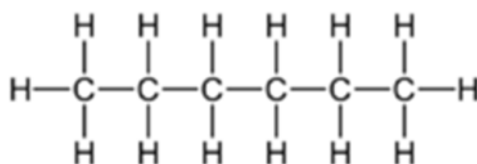


Figura 1. Molécula de hexano.

- Los enlaces entre los átomos C-C y C-H dadas por:

$$V(r) = k_e \cdot (r_e - r)^2 \quad (1)$$

donde k_e es la constante del potencial armónico, r_e la distancia de equilibrio. En el caso que nos ocupa tomarán un determinado valor para describir los enlaces C-C y C-H. Este valor vendrá dado por el campo de fuerzas elegido para hacer los cálculos.

- Los ángulos H-C-H, H-C-C y C-C-C dados por la siguiente ecuación:

$$V(\theta) = k_\theta \cdot (\theta_e - \theta)^2 \quad (2)$$

donde k_θ es la constante del potencial armónico, θ_e el ángulo de equilibrio. En el caso que nos ocupa, para el hexano, el campo de fuerzas toma un determinado valor de k_θ y θ_e para describir los ángulos H-C-H, H-C-C y C-C-C.

³ J. Harvey, *Computational Chemistry*, Oxford University Press (Oxford, 2018).

- Los ángulos diedros dados por la ecuación:

$$V(\varphi) = k_{\varphi} \cdot (1 + \cos(n \cdot \varphi - \varphi_0)) \quad (3)$$

donde k_{φ} es la constante del potencial armónico y φ_0 es el ángulo diedro en equilibrio, dados también por el campo de fuerzas elegido.

- Los términos cruzados; los cuales indican que los enlaces y ángulos no vibran de manera independiente a los enlaces o ángulos que tengan junto a ellos. Están definidos por el potencial de Urey-Bradley y dados por la ecuación:

$$V_{UB}(r_{1-3}) = k_{UB} \cdot (r_{1-3} - r_{1-3(0)})^2 \quad (4)$$

donde k_{UB} es la constante de Urey-Bradley, r_{1-3} es la distancia entre los átomos considerados como 1 y 3, y $r_{1-3(0)}$ es la distancia entre los átomos 1 - 3 optimizada del campo de fuerzas. En este caso los átomos separados por otro átomo adyacente (denominados 1-3) son los mostrados en la Figura 2.

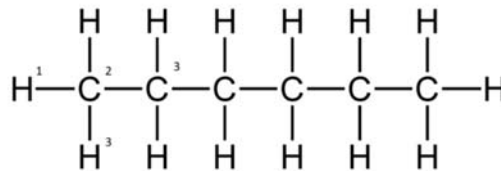


Figura 2. Molécula de hexano.

En cuanto a las interacciones no enlazantes, son:

- Electrostáticas: de Coulomb, debidas a las cargas de iones y a la polarización de los enlaces químicos y dadas por:

$$V(q_1 q_2) = \frac{1}{4\pi\epsilon_0\epsilon_R} \cdot \frac{q_1 \cdot q_2}{r_{12}} \quad (5)$$

donde q_1 y q_2 son las cargas de los átomos considerados 1 y 2, ϵ_0 es la conductividad en el vacío, ϵ_R la conductividad en el medio y r_{12} es la distancia entre los átomos 1 y 2.

- No electrostáticas, también conocidas como fuerzas débiles (van der Waals, London o dipolo-dipolo) las cuales se simulan mediante el potencial de Lennard-Jones:

$$V(r) = 4\varepsilon \cdot \left\{ \left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right\} \quad (6)$$

donde ε describe la fuerza de la interacción entre los dos átomos considerados y σ la distancia a la cual el potencial cambia de signo. Cuando $\sigma > r$, el valor del potencial es negativo y pasa a tener valores positivos cuando $\sigma < r$.³

El cálculo de los términos no enlazantes es lo que más tiempo de cálculo requiere ya que hay que considerar cada uno de los pares atómicos del sistema. La aproximación más sencilla es considerar sólo aquellos átomos que entren dentro de una esfera de un determinado radio (*cut-off*)^{1,3}. Sin embargo, este método es insuficiente para el tratamiento de las interacciones coulómicas. Por ello, para el tratamiento computacionalmente eficiente de las interacciones coulómicas a larga distancia, se usa el método denominado PME (*Particle Mesh Ewald*)³.

Hay que considerar todas las interacciones del sistema a estudiar, las enlazantes y no enlazantes entre los distintos átomos. Por lo tanto, la energía potencial total del sistema vendrá dada por:

$$\begin{aligned} V_{total} = & V_{enlazante} + V_{no-enlazante} = \\ & V_{enlace} + V_{angulo} + V_{dihedro} + V_{cruzado} \\ & + V_{Coulomb} + V_{vdWaals} \end{aligned} \quad (7)$$

El campo de fuerzas que se ha utilizado en este TFG ha sido CHARMM36⁴ (*Chemistry at Harvard Macromolecular Mechanics*). En este campo se define la topología del sistema (cómo los átomos están enlazados en la molécula) y los parámetros del hexano de las ecuaciones (1)-(6). Es un campo de fuerzas muy utilizado en la simulación de biomoléculas y el conjunto de parámetros para el hexano están basados en la parametrización de lípidos.

⁴ J. B. Klauda y cols. *J. Phys. Chem. B*, **114**, 7830 (2010).

Dinámica Molecular

La Dinámica Molecular permite simular computacionalmente la evolución de un sistema a lo largo del tiempo a nivel atómico. Está basada en el estudio de las fuerzas interatómicas mediante las ecuaciones de Newton³. Permite calcular muchas propiedades fisicoquímicas como pueden ser la entropía, la energía libre, viscosidad, temperatura, etc. Con la Dinámica Molecular se estudia cómo las moléculas se mueven a través de la SEP mediante el uso de las ecuaciones previamente resueltas mediante la Mecánica Molecular.

La geometría inicial del sistema a estudiar (hexano líquido) se ha obtenido mediante el programa GROMACS (*GRO*ningen *MA*chine for *C*hemical *S*imulation)⁵. Este es un programa de simulación de Dinámica Molecular que resuelve las ecuaciones de Newton del movimiento para sistemas de centenares a millones de partículas. Es decir, sistemas grandes, por lo que puede ser aplicado al hexano.

La geometría inicial del sistema puede ser obtenida por diversas vías. En este caso, el programa GROMACS tiene una aplicación que permite, a partir de una molécula, generar un sistema con el número de moléculas que se deseen. La posición de estas moléculas es aleatoria y aunque se tiene en cuenta que estas no solapen, es necesario realizar la minimización de la energía del sistema antes de comenzar el estudio de Dinámica Molecular.

Minimización energética

La minimización de la energía conlleva calcular el vector gradiente de la energía potencial en la dirección en que este se hace cero. Se puede calcular el gradiente de la energía potencial como el vector cuyas componentes son las derivadas de la energía potencial respecto a las diferentes coordenadas del sistema (distancias, ángulos, etc.) de acuerdo con:

⁵ B. Hess y cols. *J. Chem. Theory Comput.* **4**, 435 (2008).

$$\nabla V = \left(\frac{dV}{dr_1}, \frac{dV}{dr_2}, \frac{dV}{d\theta_1}, \dots \right) \quad (8)$$

Por ejemplo, con las expresiones clásicas del campo de fuerzas es relativamente sencillo hallar la energía potencial porque la derivada de la energía con respecto a las coordenadas es un cálculo rápido de realizar, tal y como se muestra a continuación tomando como ejemplo la expresión de la energía potencial de un enlace (ver Figura 3):

$$V(r) = k_e \cdot (r_e - r)^2 \quad (9)$$

$$\frac{dV(r)}{dr} = -2k_e \cdot (r_e - r) \quad (10)$$

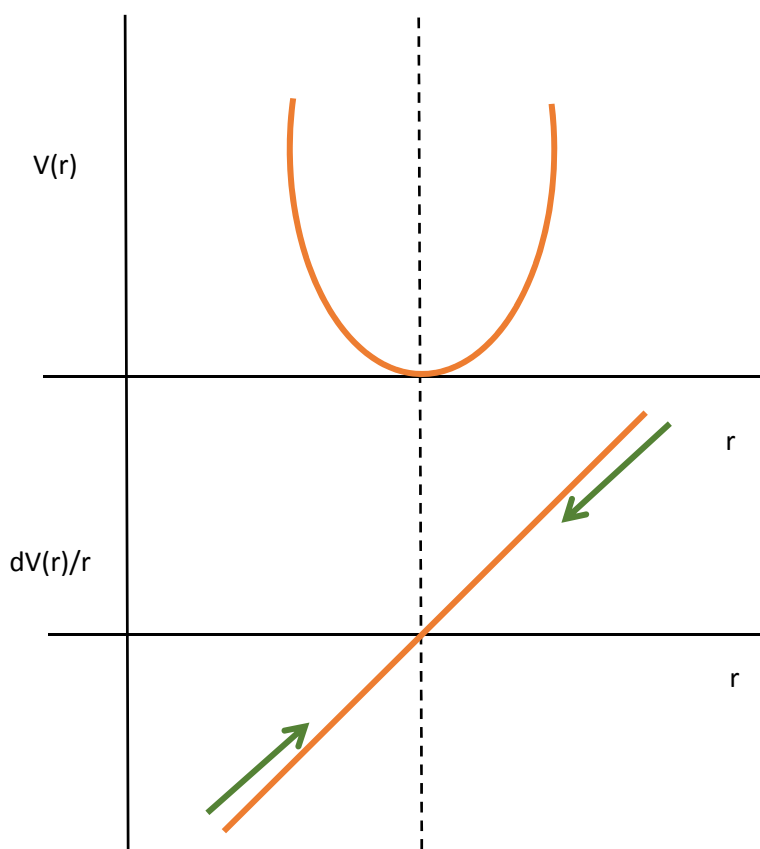


Figura 3. Relación simplificada entre dos átomos y su gradiente.

El signo de cada componente del gradiente indica el sentido a modificar (ampliar o reducir) la coordenada; si se tiene una distancia muy pequeña con respecto a la de equilibrio el valor del gradiente en ese punto se hace negativo así, se tendrá que

augmentar r para alcanzar la posición de equilibrio, y si se tiene una distancia muy grande el valor es positivo y se sigue el sentido contrario.

En el proceso de minimización es necesario definir un paso ϵ que indica el tanto por uno a variar la coordenada en cuestión. El valor elegido en este trabajo fue 0,01 nm. En cuanto al valor final del gradiente de la energía, numéricamente es muy difícil llegar a un valor de gradiente cero, por lo que se utiliza el valor de 1000 kJ mol⁻¹ nm⁻¹. Estos valores son utilizados por defecto por GROMACS⁶, ya que es suficiente para evitar contactos entre las moléculas y átomos en la geometría inicial del sistema a estudiar.

El número total de pasos utilizado en este TFG en la minimización fue 50.000. El proceso de minimización cesa cuando llega al valor del gradiente mínimo propuesto que puede alcanzarse antes de los 50.000. En este TFG no fue necesario llegar en ninguno de los cálculos al máximo total de 50.000.

Otro efecto a considerar es el debido a la temperatura. En la Mecánica Clásica el efecto de la temperatura influye en la energía cinética y esta, a su vez, en la velocidad de las partículas. La relación de la velocidad de las partículas del sistema con la temperatura viene dada por la Distribución de Maxwell². El primer paso en una simulación de Dinámica Molecular es el calentamiento, el cual debe ser paulatino. Para ello, a los átomos de la estructura minimizada se les asigna una velocidad conforme con la Distribución de Maxwell dada por:

$$\frac{dN_v}{N} = \left(\frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} e^{\frac{-mv^2}{2k_B T}} \cdot 4\pi v^2 dv$$

$$\frac{dN_{v_x}}{N} = \left(\frac{m}{2\pi k_B T} \right)^{\frac{1}{2}} e^{\frac{-mv_x^2}{2k_B T}} \cdot 4\pi v^2 dv_x$$
(11)

donde m es la masa de la partícula, k_B es la constante de Boltzmann, T es la temperatura absoluta del sistema y v es la velocidad de las partículas. En las figuras 4 y 5 se representa la distribución de Maxwell gráficamente para la molécula de Nitrógeno. Aunque no está

⁶ Manual GROMACS: <http://manual.gromacs.org/documentation/> (última fecha de acceso 19 de julio de 2018).

relacionado con el sistema estudiado, da una idea de la forma de las distribuciones y el efecto de la temperatura sobre estas.

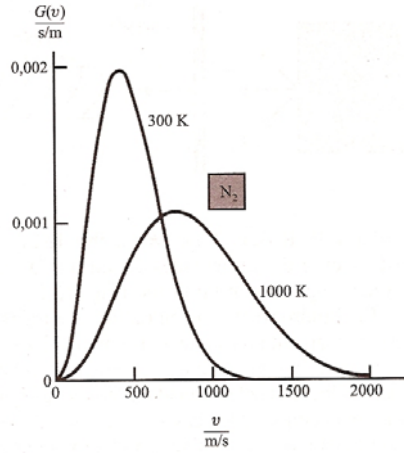


Figura 4: Ley de distribución de Maxwell.

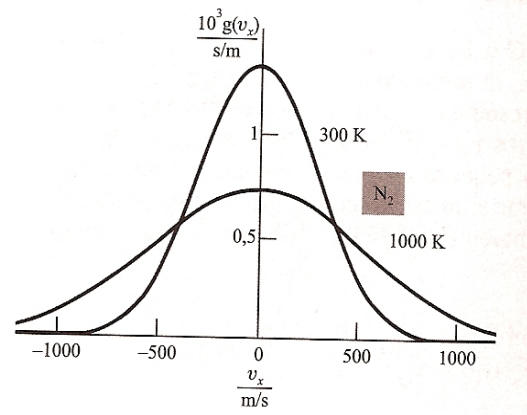


Figura 5: Ley de distribución de Maxwell.

Conociendo la posición, la velocidad y la aceleración de todas las partículas del sistema se puede conocer la trayectoria, es decir, las distintas posiciones que adquieren las partículas a lo largo del tiempo. En los estudios dinámicos la trayectoria recorre el espacio-fase del sistema. Es decir, todas las posiciones y momentos lineales de las partículas del sistema, tal y como se observa en la siguiente ecuación:

$$\begin{aligned}\vec{r}(t + \Delta t) &= \vec{r}(t) + \frac{\vec{p}(t)}{m} \Delta t = \vec{r}(t) + \vec{v}(t) \cdot \Delta t \\ \vec{v}(t + \Delta t) &= \vec{v}(t) + \vec{a}(t) \cdot \Delta t \\ \vec{a}(t) &= \frac{\vec{F}(t)}{m}\end{aligned}\tag{12}$$

donde m es la masa de cada partícula, $\vec{r}(t)$, $\vec{p}(t)$, $\vec{v}(t)$ y $\vec{a}(t)$ son la posición, momento lineal, velocidad y aceleración de estas a un tiempo t . La fuerza, $\vec{F}(t)$, que actúa sobre las partículas en cada instante se puede hallar mediante el campo de fuerzas conocidas las velocidades, de acuerdo con:

$$\vec{F}(t) = \nabla V(\vec{r}) = - \frac{dV(\vec{r})}{d\vec{r}(t)}\tag{13}$$

El programa de simulación, mediante el denominado propagador,⁶ calcula estas magnitudes a diferentes tiempos comenzando desde la posición inicial del sistema ($t=0$).

El incremento de tiempo de la simulación, Δt , no debe ser muy pequeño dado que en este caso el número de cálculos a realizar aumenta de forma desorbitada y la simulación tiene un alto coste computacional. Tampoco tiene que ser muy grande porque en este caso el sistema colapsa. En este TFG el tiempo escogido para las simulaciones realizadas fue 1 fs, menor que el comúnmente escogido (2 fs)⁶, dado que al querer estudiar una propiedad dependiente de la posición se ha buscado tener más precisión en las simulaciones a lo largo del tiempo.

La simulación debe realizarse para un sistema en continuo. Es decir, conforme se aumenta la velocidad del sistema, los átomos difundirán hacia el vacío. Para evitar que esto ocurra se utilizan las denominadas condiciones periódicas de contorno (*Periodic Boundary Conditions*, PBC)^{3,6}. En este modelo la simulación se extiende de forma periódica en las tres direcciones del espacio. Las PBC permiten que la simulación sea en continuo; las interacciones no enlazantes se consideran en las imágenes periódicas adyacentes. Por ello, se elige el tamaño de la caja de manera que un átomo no vea la imagen de sí mismo (Figura 6) puesto que esto produciría errores de cálculo.

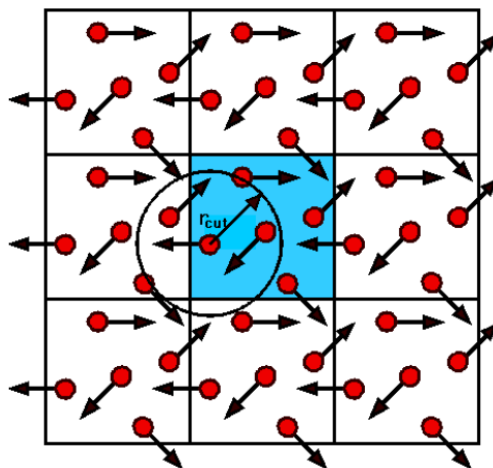


Figura 6: Esquema de las condiciones periódicas de contorno.⁷

Durante la dinámica es importante ir comprobando que la temperatura es la deseada. Esto se puede realizar gracias a las propiedades de la distribución de Maxwell, y se conoce comúnmente como termostato que se muestra en la ecuación:

⁷ http://wiki.cs.umt.edu/classes/cs477/index.php/Distance_Matrix (última fecha de acceso 19 de julio de 2018).

$$\begin{aligned}\langle v^2 \rangle &= \frac{3RT}{M} \\ T(t) &= \frac{1}{3R} \frac{\sum_{i=1}^n m_i v_i^2(t)}{N}\end{aligned}\tag{14}$$

donde v_i es la velocidad de la partícula i , m_i su masa, N el número total de átomos del sistema y M la masa total de las partículas.

La estructura inicial de partida (una vez minimizada su energía) se somete a un proceso de calentamiento asignando las velocidades según la Distribución de Maxwell hasta obtener la temperatura deseada. Tras esto, se deja evolucionar el sistema a lo largo del tiempo. En esta fase se controla que la temperatura de la simulación se mantiene constante a través del termostato, reasignando las velocidades si hubiera desviación⁶. Es la denominada etapa de equilibrado.

Una vez equilibrado el sistema comienza la siguiente etapa, la etapa de producción. Es la etapa más larga y en la que se van guardando los datos de la simulación (estructura, velocidad y propiedades termodinámicas). Una vez finalizada esta etapa se analiza la propiedad fisicoquímica que interese a partir de los datos almacenados durante la simulación.

El tipo de propiedad a analizar viene definido por las condiciones de simulación⁶. En este TFG, en el que se ha querido determinar el coeficiente de difusión de un líquido, se ha escogido el colectivo canónico o NVT en el que el número de partículas (N), el volumen (V) y la temperatura (T) son constantes.

Cálculo del coeficiente de difusión y de la viscosidad

El coeficiente de difusión es una propiedad correlacionada con el tiempo. Es decir, su valor a un determinado tiempo influye en el valor de sí misma a un tiempo mayor. Para el estudio de este tipo de propiedades se utilizan las denominadas funciones de autocorrelación. Así, a partir de la ecuación de Einstein-Smoluchovski, el coeficiente de

difusión de una partícula calculada a través del análisis de la trayectoria de N partículas es⁸:

$$D = \frac{1}{6N} \lim_{t \rightarrow \infty} \frac{d}{dt} \left\langle \sum_{i=1}^N [\vec{r}_i(t) - \vec{r}(0)]^2 \right\rangle \quad (15)$$

donde $[\vec{r}_i(t) - \vec{r}(0)]^2$ es el desplazamiento cuadrático de la posición de la molécula i y N el número total de moléculas. Las posiciones a tiempo cero son las posiciones de las moléculas en la estructura inicial y el promedio del desplazamiento cuadrático de la posición para todas las moléculas es el Desplazamiento Cuadrático Medio (DCM). En la Figura 7 se muestra un ejemplo de la variación del DCM con respecto del tiempo.

Naturalmente el tiempo de simulación es limitado ($t \neq \infty$), así el coeficiente de difusión se calcula a partir de la pendiente de la representación del DCM frente al tiempo dividida por 6, de acuerdo con la ecuación (15).

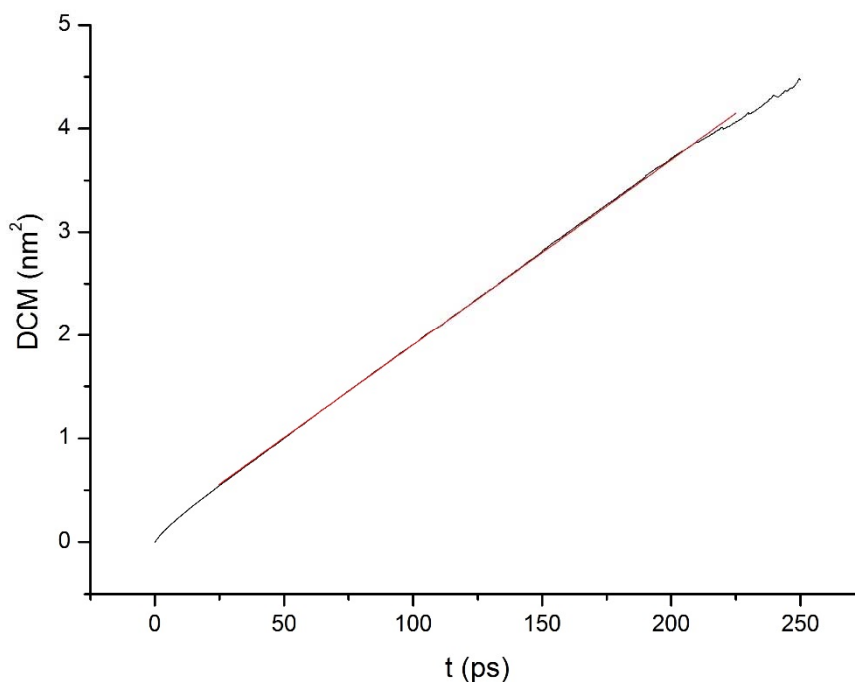


Figura 7: Representación de DCM frente al tiempo.

⁸ L. H. Eckler y M.J. Nee, *J. Chemical Ed.*, **93**, 917 (2016).

Otra propiedad de transporte interesante a calcular con este tipo de simulaciones es la viscosidad del líquido. Para determinar la viscosidad se usan las funciones de correlación de estrés basadas en las simulaciones a diferentes presiones^{6,9}. Este tratamiento está fuera de los objetivos de este Trabajo Fin de Grado, pero una alternativa sencilla a este procedimiento es el uso de la ecuación de Stokes-Einstein².

El coeficiente de difusión para partículas consideradas esféricas se calcula mediante la ecuación de Stokes-Einstein:

$$D = \frac{k_B \cdot T}{6 \cdot \pi \cdot \eta \cdot r} \quad (16)$$

donde η es la viscosidad del medio, r el tamaño de la partícula y T la temperatura. Se puede observar que cuanto mayor es la viscosidad del medio, menor es el coeficiente de difusión de la partícula. Esto es debido a que la viscosidad hace referencia a la fricción que el medio opone al movimiento de la partícula.

La ecuación (16) se puede aplicar a la difusión en disolución siempre que el radio de la partícula que difunde sea mayor que el radio de las partículas del líquido.¹⁰ Sin embargo, para la difusión de partículas en líquidos con tamaños de partícula similares, la ecuación (16) se modifica a:

$$D = \frac{k_B \cdot T}{4 \cdot \pi \cdot \eta \cdot r} \quad (17)$$

La ecuación (17) también puede aplicarse al fenómeno de la autodifusión, consistente en la difusión de moléculas de disolvente en el propio disolvente. El hecho de modificar el denominador de la ecuación de Stokes-Einstein para problemas de difusión de partículas de tamaño similar, y para el fenómeno de autodifusión, ha quedado demostrado a partir de datos experimentales¹⁰, así como mediante simulaciones aplicando la Dinámica Molecular para sistemas formados por un disolvente y partículas que difunden en él de igual tamaño¹¹⁻¹².

⁹ B. J. Palmer, *Phys. Rev. E*, **49**, 359 (1994).

¹⁰ T. Engel y P. Reid, *Química Física*, Pearson Education (Madrid, 2006).

¹¹ F. Ould-Kaddour y D. Levesque; *Phys. Rev. E*, **63**, 11295 (2001).

¹² S.H. Lee, *Bull. Korean Chem. Soc.* **33**, 735 (2012).

Al realizar el cálculo del coeficiente de difusión de acuerdo con la ecuación (15), nos encontramos con el problema de trabajar con cajas de tamaño limitado. Al ser la difusión una propiedad que depende del desplazamiento de las moléculas, el uso de las propiedades periódicas de contorno (Figura 6) da problemas ya que el movimiento de las moléculas no es real, dado que cuando una molécula sale por un lado de la caja entra por el otro lado. Idealmente la simulación se debería hacer en una caja de tamaño infinito o muy grande pero esto último conlleva un esfuerzo computacional enorme. La alternativa aplicada en este trabajo para evitar este problema es el cálculo del coeficiente de difusión con tamaños de caja distintos. De esta manera, la ordenada en el origen de la representación de los valores del coeficiente de difusión D (calculados a la misma temperatura) frente a la inversa de la longitud del lado de la caja será el valor de D corregido, ya que en la ordenada en el origen, el lado de la caja es idealmente infinito¹³. Una representación gráfica de este tipo de cálculo se presenta en la Figura 8.

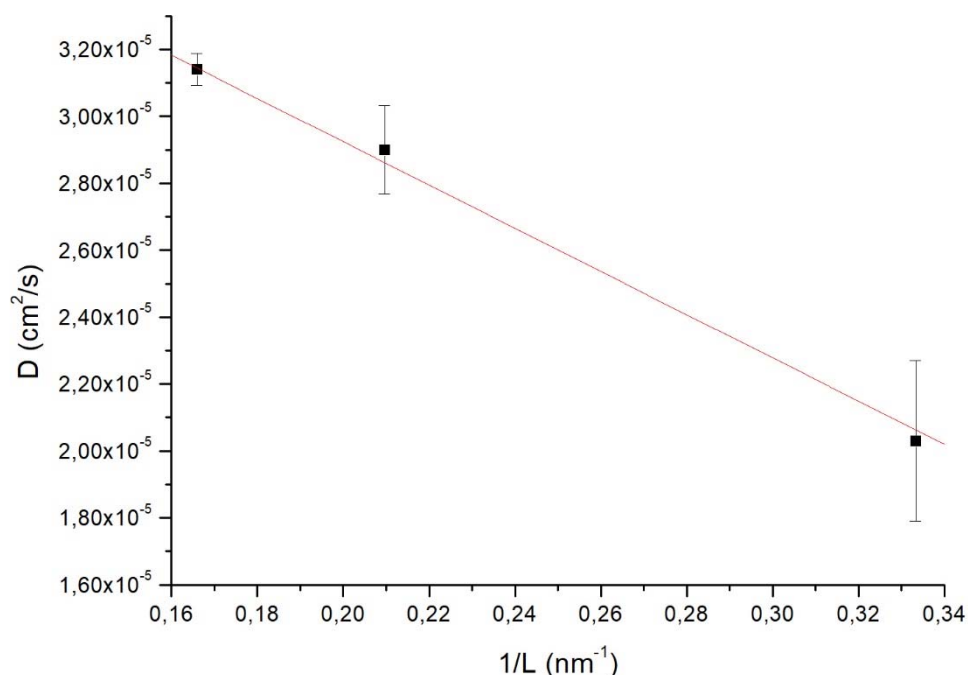


Figura 8: Representación del coeficiente de difusión calculado a 320 K frente a $1/L$ para las tres cajas estudiadas (caja pequeña $l= 3 \text{ nm}$, caja mediana $l= 4,17 \text{ nm}$ y caja grande $l=6 \text{ nm}$).

¹³ I-C. Yeh y G. J. Hummer, *Phys. Chem. B*, **108**, 15873(2004).

PROCEDIMIENTO COMPUTACIONAL

El proceso computacional comprende los siguientes pasos:

Primero, se lleva a cabo la creación de la caja inicial. Se parte de una molécula de hexano y se crea una caja con el número de moléculas deseado en posición aleatoria (ver detalles más adelante). En este trabajo se crearon tres cajas de número total de moléculas de hexano 128, 512 y 1024 para el estudio del efecto del tamaño de la caja sobre el coeficiente de difusión.

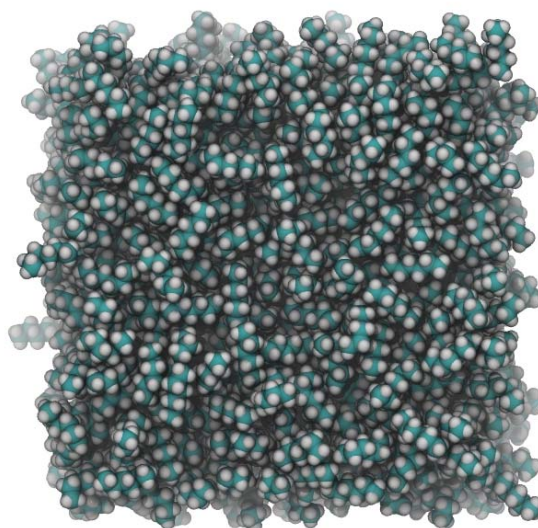


Figura 9. Caja cúbica de 1024 moléculas de hexano.

Después, se realiza una minimización energética con el método *Steepest Descent*⁶ seguido de una dinámica de calentamiento de 400 ps hasta la temperatura deseada y en condiciones NPT a una presión de 1 bar (baróstato de Berendsen⁶). En estas condiciones el sistema evoluciona ajustando su volumen según la densidad del sistema, si el campo de fuerzas es apropiado reproducirá el valor experimental.

Tras la dinámica de calentamiento y equilibrado, se realiza la dinámica de producción en condiciones NVT durante 250 ps con un Δt de 1 fs, analizando la trayectoria resultante para el cálculo del coeficiente de difusión según la ecuación (17). Este último paso se repite tres veces variando la asignación de velocidades iniciales con el fin de obtener tres trayectorias de producción en condiciones iniciales distintas, siendo el coeficiente

de difusión para una temperatura y tamaño de caja determinado la media de estos valores. Se calcula su error a un 95% de confianza según la distribución t de Student¹⁴.

En la Figura 10 se muestran los pasos que se han seguido desde el estudio de la dinámica hasta la obtención de las propiedades finales del sistema (D y η).

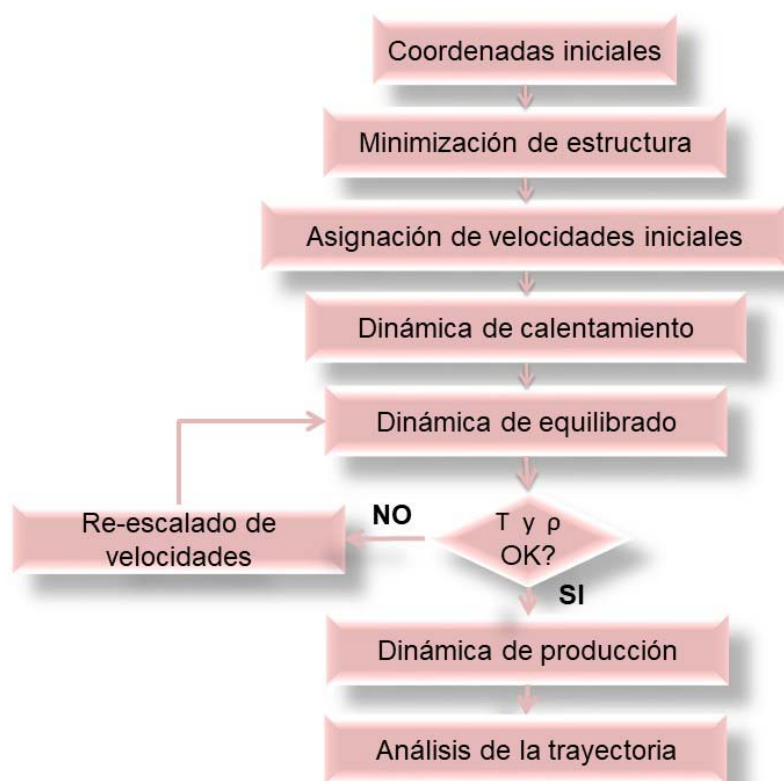


Figura 10: Esquema de trabajo.

En este Trabajo Fin de Grado se han realizado las simulaciones para tres tamaños de caja ($l = 3$ nm, 4,17 nm y 6 nm) con 128, 512 y 1024 moléculas de hexano a las temperaturas de trabajo 260, 280, 300, 320 y 340 K. Cada dinámica de producción se repitió tres veces a cada temperatura de trabajo con el fin de obtener tres réplicas para cada caja. Ello implica un total de 45 trayectorias de producción.

Los cálculos se han realizado con el programa GROMACS versión 5.3. utilizando la distribución Ubuntu16.04 (64 bits)¹⁵. Para ello, se crea una máquina virtual (VirtualBox

¹⁴ J. N. Miller y J. C. Miller, *Estadística y quimiometría para química analítica*, Prentice Hall (Madrid, 2002).

¹⁵ <https://www.ubuntu.com/> (última fecha de acceso 19 de julio de 2018).

versión 5.2.6 r120293(Qt5.6.2))¹⁶ sobre el ordenador de trabajo en la que se puede instalar un Sistema Operativo funcional que no afecta al ordenador de trabajo.

Una parte de los cálculos se ha realizado en los ordenadores de las aulas informáticas de la Fase IV de la Facultad de Ciencia y Tecnología de la Universidad de La Rioja provistas de la máquina virtual. Estos ordenadores cuentan con un procesador Intel(R) Core (TM) i5-6600 CPU 3.30 GHz y 8.00 Gb de memoria RAM (utilizables 7.88 Gb). Los cálculos realizados en estos ordenadores se ejecutan en paralelo usando 2 núcleos del procesador.

Para los cálculos más costosos desde el punto de vista computacional se utilizó una máquina de cálculo del Laboratorio Computacional de Química Física provista de un procesador Intel(R) Xeon(R) E5620 de 2.40 GHz y 12 GB de RAM con 8 núcleos trabajando en paralelo.

A continuación se van a explicar las diferentes órdenes que se han introducido en la terminal para realizar todos los cálculos. En el ejemplo mostrado se explican los comandos para la caja computacionalmente más costosa (N=1024 moléculas de hexano).

La geometría inicial de la molécula de hexano fue generada con el programa Avogadro¹⁷ en formato PDB. Es necesario modificar el nombre y definición del residuo para que este sea identificado por el programa GROMACS. Una vez modificado este fichero se crea la topología del mismo:

```
gmx pdb2gmx -f hexano.pdb -o hexano.gro -p topol.top
```

Esta orden crea el fichero de la geometría en formato del programa GROMACS hexano.gro y la topología, topol.top, donde se encuentran los parámetros del campo de fuerzas específicos del hexano y su conectividad. El campo de fuerzas es elegido entre los ofrecidos por el programa una vez introducida la orden.

Para crear la caja con un número determinado de moléculas se usa la orden.

```
gmx insert-molecules -f hexano.gro -ci hexano.gro -o box1024.gro -box 15 15 15 -nmol 1023
```

¹⁶ <https://www.virtualbox.org/> (última fecha de acceso 19 de julio de 2018).

¹⁷ <https://avogadro.cc/> (última fecha de acceso 19 de julio de 2018).

Esta orden toma 1023 (-nmol 1023) veces la estructura del hexano (-ci hexano.gro) la introduce aleatoriamente junto con la del hexano ya existente (-f hexano.gro) y la introduce en una caja cúbica de 15 nm de lado (-box 15 15 15) en el fichero box1024.gro. El tamaño de la caja inicial es mayor que la obtenida tras la fase de equilibrado. Esto es debido a que el programa introduce aleatoriamente las moléculas de hexano y hay que definir una caja inicial muy grande para evitar que el programa dé error o introduzca un número menor de moléculas.

Una vez obtenida una caja inicial, el siguiente proceso es la minimización energética del sistema para evitar contactos entre moléculas que después den problemas al introducir la temperatura. Antes de este paso hay que modificar el fichero de topología del sistema ya que anteriormente este estaba definido para 1 molécula y ahora hay que indicar que hay 1024 moléculas en el sistema.

```
gmx grompp -f em.mdp -c box1024.gro -p topol1024.top -o em.tpr
```

```
gmx mdrun -deffnm em -v
```

Con la primera orden el programa lee los parámetros del proceso de minimización del fichero em.mdp, lee la estructura inicial, box1024.gro, y la topología del sistema, topol1024.top. Esta orden genera un archivo binario, em.tpr, con toda la información anterior que es leído y ejecutado mediante la siguiente orden por el programa que realmente es el encargado de ejecutar los procesos requeridos y que genera todos los ficheros de estructura final (em.gro), energía (em.edr) y trayectoria (em.trr) con el nombre por defecto em (-deffnm).

Esta estructura optimizada se utiliza en la dinámica de calentamiento, que a la vez es la de equilibrado. Esta fase se realiza en condiciones NPT para que el sistema, una vez alcanzada la temperatura deseada, adopte el volumen que le corresponde a una presión de 1 bar.

```
gmx grompp -f npt.mdp -c em.gro -p topol1024.top -o npt.tpr
```

```
gmx mdrun -deffnm npt -v
```

En el fichero npt.mdp se indican las condiciones del cálculo como la presión de 1 bar o la temperatura que variará según el experimento (260, 280, 300, 320 y 340 K). Una variable importante a introducir es el tiempo de simulación, en este caso se comprobó

que con 250 ps era suficiente para que el sistema adoptara un volumen constante a lo largo del tiempo final de simulación.

La validez del campo de fuerzas así como de los parámetros de simulación utilizados se realiza mediante el análisis de la densidad de la mezcla que se comprobó con una simulación a 300 K. Al estar el sistema previamente equilibrado el resultado de este cálculo debe coincidir con el dato experimental. Este análisis se realiza mediante la orden:

```
gmx energy -f nvt.edr
```

El menú de respuesta del programa de análisis *energy* permite elegir muchas propiedades a analizar como la temperatura, energía cinética, energía potencial, energía total y presión. En este caso se eligió la densidad ofreciendo un valor de $672 \pm 5,00$ kg/m³, dato con una desviación aceptable (2,0 %) con respecto al resultado experimental¹⁸ de 660,6 kg/m³.

La etapa de producción se llevó a cabo a volumen constante, leyendo como estructura inicial, la final de la etapa anterior que corresponde al sistema ya equilibrado.

```
gmx grompp -f prod.mdp -c npt.gro -p topol1024.top -o prod.tpr
```

```
gmx mdrun -deffnm prod -v
```

Este cálculo se repite tres veces modificando el número pseudo-aleatorio que asigna las velocidades iniciales a las moléculas, de esta manera se obtienen tres trayectorias distintas que dan lugar a tres valores del coeficiente de difusión cuya media se utilizará en el cálculo del coeficiente de difusión corregido. Este se calcula analizando la trayectoria de la simulación

```
gmx msd -f prod.trr -s prod.tpr
```

El programa de análisis *msd* realiza el cálculo según la Figura 7 y la ecuación (15) dando un valor del coeficiente de difusión en pantalla directamente.

La descripción de los archivos em.mdp, npt.mdp y prod.mdp se muestra en el Anexo 1.

¹⁸ W.Haynes, *Handbook of Chemistry and Physics*, 91st. ed. (Boca Ratón, 2011).

RESULTADOS

En las figuras 11-15 se presentan los resultados obtenidos para los valores de D a cada temperatura estudiada para los diferentes tamaños de caja investigados. En el Anexo 2 se presentan las tablas con todos los valores representados en las figuras. Además, en el Anexo 3 se muestran las figuras con los valores de los datos de los ajustes lineales obtenidos.

Se observa que existe una cierta dispersión en los resultados del coeficiente de difusión calculados considerando el mismo tamaño de caja. Además se observa que todas las pendientes ajustadas son negativas lo que indica que cajas de simulación pequeñas subestiman el valor de D .

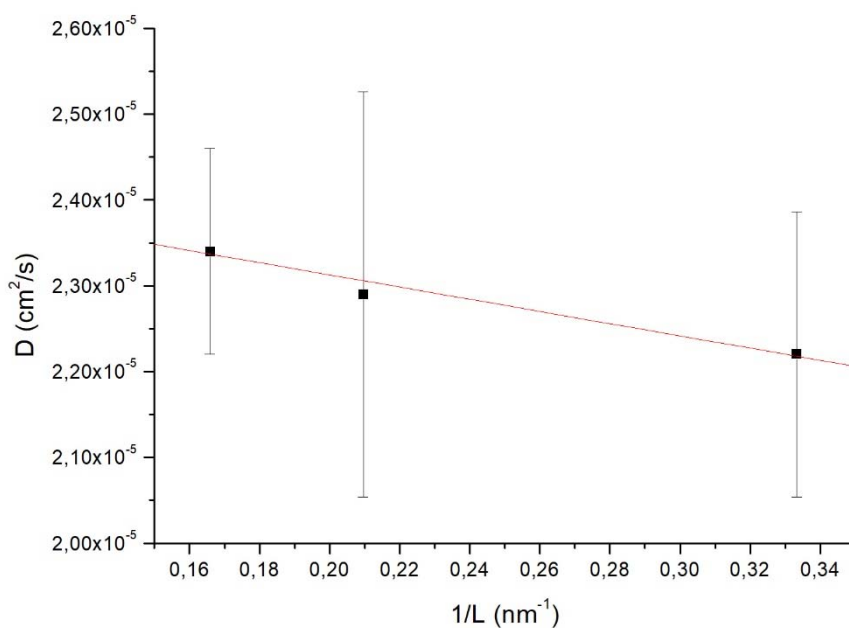


Figura 11: Representación del coeficiente de difusión respecto al inverso de la longitud a 260K con las tres cajas empleadas en el análisis.

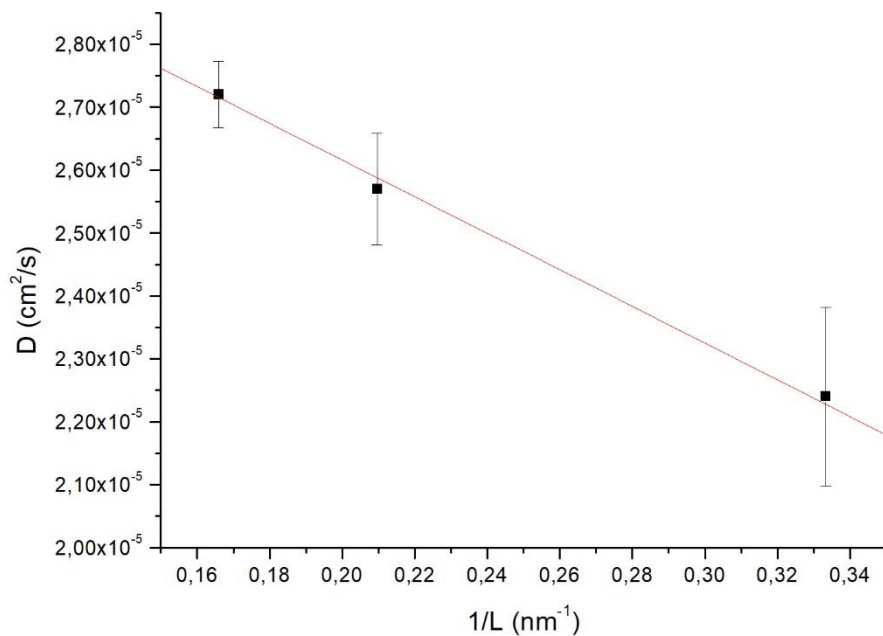


Figura 12: Representación del coeficiente de difusión respecto al inverso de la longitud a 280K con las tres cajas empleadas en el análisis.

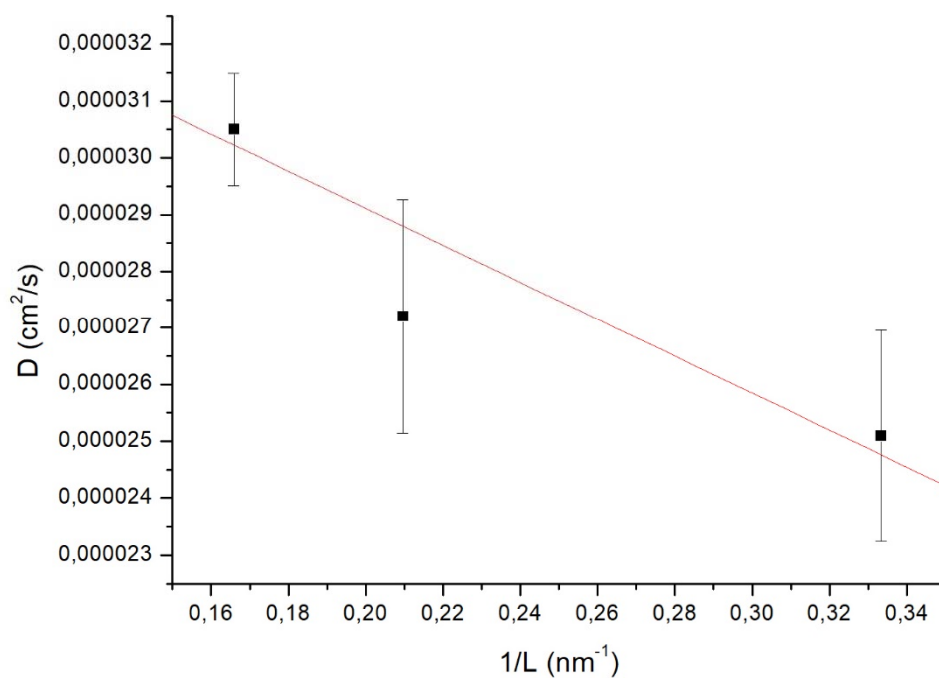


Figura 13: Representación del coeficiente de difusión respecto al inverso de la longitud a 300K con las tres cajas empleadas en el análisis.

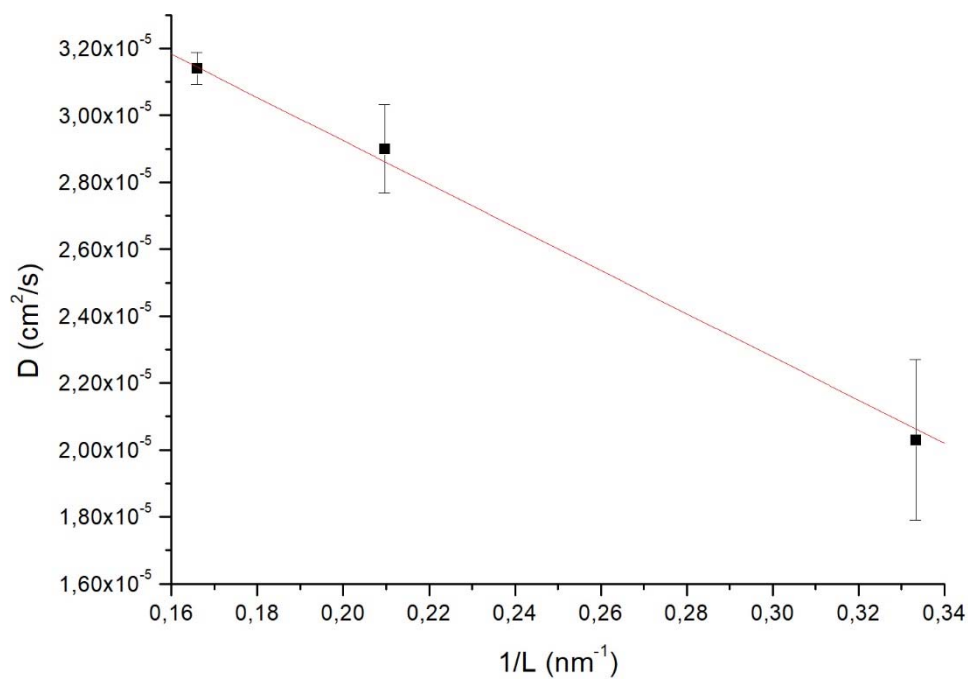


Figura 14: Representación del coeficiente de difusión respecto al inverso de la longitud a 320K con las tres cajas empleadas en el análisis.

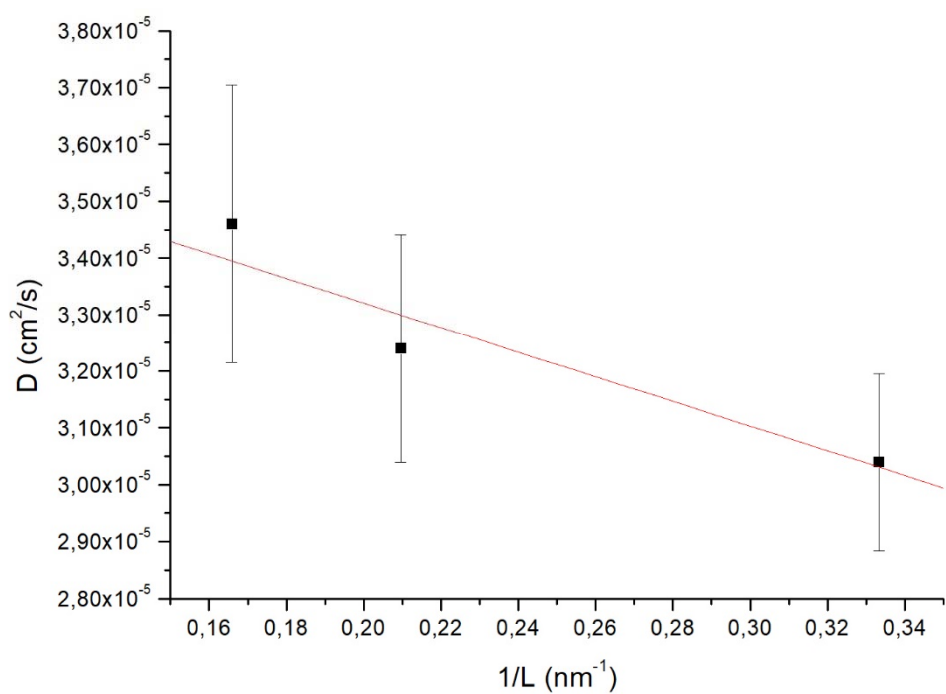


Figura 15: Representación del coeficiente de difusión respecto al inverso de la longitud a 340K con las tres cajas empleadas en el análisis.

A partir de las figuras 11-15 se calcula el valor de los coeficientes de difusión corregidos tal y como se ha comentado en el apartado de Introducción Teórica (Figura 8), ya que es necesario aplicar esta corrección debida a que el tamaño de la caja de simulación es finito. Estos valores corregidos se muestran en la Tabla 1 junto con el cálculo de su error. Tal y como se observa, conforme aumenta la temperatura el valor del coeficiente de difusión aumenta, a excepción del cálculo a 320K.

Tabla 1: Coeficientes de difusión y errores de difusión a las temperaturas analizadas.

T (K)	D (cm ² /s)
260	$3,25 \cdot 10^{-5} \pm 6,84 \cdot 10^{-6}$
280	$3,18 \cdot 10^{-5} \pm 3,91 \cdot 10^{-7}$
300	$3,44 \cdot 10^{-5} \pm 2,99 \cdot 10^{-6}$
320	$4,28 \cdot 10^{-5} \pm 7,87 \cdot 10^{-7}$
340	$3,79 \cdot 10^{-5} \pm 1,71 \cdot 10^{-6}$

En la Tabla 2 se muestran los valores de los coeficientes de difusión obtenidos experimentalmente^{19,20,21,22,23,24} a diversas temperaturas de trabajo. Una representación gráfica de los resultados experimentales y los obtenidos en las simulaciones se muestra en la Figura 15.

¹⁹ G.M.Panchencov y V.V.Erchenkov, *R.J.Phys.Chem* **36**,455 (1962).

²⁰ Y.Yamaguchi y cols., *Chem. Soc. Japan* **63**, 2154 (1990).

²¹ D. C. Douglas y D.W. McCall, *J.Phys. Chem*, **62**, 1102 (1958).

²² D. C. Douglas y D.W. McCall, *Phys. Fluids*, **2**, 1102 (1959).

²³ M.A.Awan y J.H.Dymond, *I.J.Thermophys*, **22**, 679.

²⁴ C.D.Agostino y cols., *Chem.Eng.Sci*, **66**, 3898 (2011).

Tabla 2: Datos experimentales del coeficiente de difusión a diferentes temperaturas.

T (K)	D (cm ² /s)
193,2	$6,20 \cdot 10^{-5}$
213,2	$1,05 \cdot 10^{-6}$
233,2	$1,63 \cdot 10^{-5}$
253,2	$2,25 \cdot 10^{-6}$
273,2	$3,00 \cdot 10^{-5}$
293,2	$3,85 \cdot 10^{-6}$
294,15	$3,93 \cdot 10^{-6}$
298,1	$4,44 \cdot 10^{-6}$
298,15	$4,21 \cdot 10^{-5}$
313,2	$4,86 \cdot 10^{-6}$
323,1	$5,96 \cdot 10^{-6}$
333,2	$6,00 \cdot 10^{-6}$

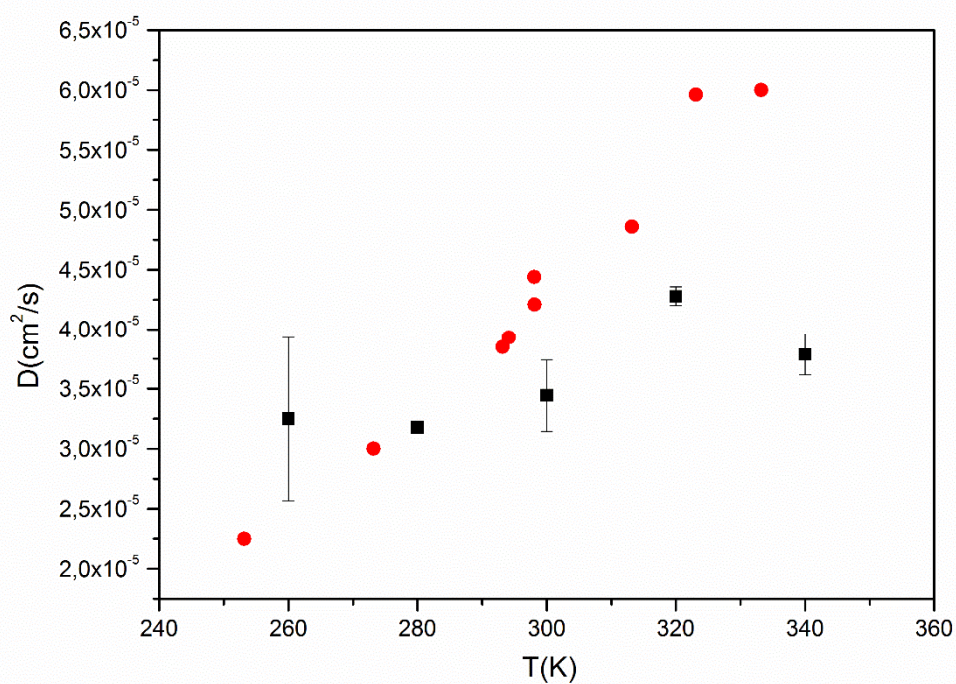


Figura 16: Representación del coeficiente de difusión respecto a la temperatura. Los puntos rojos representan los datos experimentales¹⁹⁻²⁴. Los cuadrados negros son los valores obtenidos mediante cálculos computacionales.

Como se observa en la Figura 16, la tendencia de los valores experimentales muestra un incremento en el valor del coeficiente de difusión D con la temperatura. Esta tendencia también se observa en los resultados calculados, a excepción de los valores a 260 y 340 K. Hay que tener en cuenta que la parametrización del campo de fuerzas aplicado se realiza para describir propiedades de interés biológico alrededor de 300 K, de ahí que los datos entre 280 y 320 K sean los que dan lugar a la mejor comparación.

Una posible mejora de los resultados obtenidos para D consistiría en ampliar y mejorar el estudio de la dependencia del coeficiente de difusión con el tamaño de las cajas. Es decir, realizar las simulaciones para un mayor número de cajas y con mayor número de moléculas de hexano. Sin embargo, esto fue imposible de realizar debido al alto coste computacional que conlleva. También, la ampliación del número de réplicas por sistema mejoraría la calidad de la comparación ya que se ha visto que a pesar de que el cálculo está convergido y que la trayectoria es suficientemente larga (Figura 7), el cambio en la asignación de las velocidades iniciales conduce a variaciones en D .

Una vez obtenidos los coeficientes de difusión para las diferentes temperaturas de trabajo, se puede calcular la viscosidad del sistema hexano según la ecuación de Stokes-Einstein (ecuación 17). Además, se han calculado su incertidumbre utilizando la expresión de la propagación de errores¹⁴:

$$y = f(x)$$

$$\Delta y^2 = \left| \frac{dy}{dx} \right|^2 \cdot \Delta x^2 \quad (18)$$

$$\Delta \eta = \frac{k_B \cdot T}{4\pi \cdot r \cdot D^2} \cdot \Delta D$$

Para el cálculo de la viscosidad y su incertidumbre, es necesario conocer el radio de la molécula de hexano, r , suponiéndola como partícula esférica.

Así, para el sistema de $N=128$ moléculas de hexano, el volumen de la caja cúbica de lado $l=3,0$ nm será $27,0$ nm³, de manera que el volumen de una molécula de hexano considerada esférica y que ocupa todo el espacio, está dado por:

$$V_{hex} = \frac{V_T}{N} = \frac{27}{128} = 0,211 \text{ nm}^3 \quad (19)$$

Para el sistema de $N=512$ moléculas de hexano, el volumen de la caja cúbica de lado $l=4,17$ nm es $108,53$ nm³, de manera que, en este caso, el volumen de una molécula de hexano esférica es:

$$V_{hex} = \frac{V_T}{N} = \frac{108,53}{512} = 0,212 \text{ nm}^3 \quad (20)$$

Finalmente, para el sistema de $N=1024$ moléculas de hexano, el volumen de la caja cúbica de lado $l=6,0$ nm, es $216,00$ nm³, por lo que el volumen de una molécula de hexano será:

$$V_{hex} = \frac{V_T}{N} = \frac{216,00}{1024} = 0,211 \text{ nm}^3 \quad (21)$$

De manera que, considerando los valores calculados para el volumen de la molécula de hexano en las ecuaciones (19)-(21), se ha tomado el volumen del hexano igual a $0,211$ nm³. Por lo tanto, considerando la molécula de hexano como una esfera, se tiene que:

$$V_{hex} = \frac{4}{3} \cdot \pi \cdot r^3 \quad (22)$$

$$r = \sqrt[3]{\frac{3 \cdot v_{hex}}{4\pi}} = \sqrt[3]{\frac{3 \cdot 0.211}{4\pi}} = 0,37 \text{ nm}$$

En las tablas 3 y 4 se muestran los valores de la viscosidad calculados a las diferentes temperaturas de trabajo junto con su error (Tabla 3), y los valores determinados experimentalmente^{25,26} (Tabla 4).

Tabla 3: Valores de viscosidad y sus errores calculados a las temperaturas de trabajo

T (K)	η (mPa·s)
260	$2,38 \cdot 10^{-1} \pm 5,00 \cdot 10^{-2}$
280	$2,62 \cdot 10^{-1} \pm 3,22 \cdot 10^{-3}$
300	$2,59 \cdot 10^{-1} \pm 2,25 \cdot 10^{-2}$
320	$2,22 \cdot 10^{-1} \pm 4,09 \cdot 10^{-3}$
340	$2,66 \cdot 10^{-1} \pm 1,20 \cdot 10^{-2}$

²⁵ Y. M. Naziev y cols, *Izv. Vyssh. Uchebn. Zaved. Neft. Gaz*, **15**, 65 (1972).

²⁶ S.I.Mekhtiev y cols, *Izv. Vyssh. Uchebn. Zaved. Neft. Gaz*, **3**, 64 (1975).

Tabla 4: Datos experimentales de la viscosidad a diferentes temperaturas^{25, 26}.

T (K)	η (mPa·s)
223	$7,439 \cdot 10^{-1}$
233	$6,398 \cdot 10^{-1}$
243	$5,525 \cdot 10^{-1}$
253	$4,788 \cdot 10^{-1}$
263	$4,284 \cdot 10^{-1}$
273	$3,827 \cdot 10^{-1}$
283	$3,442 \cdot 10^{-1}$
293	$3,131 \cdot 10^{-1}$
303	$2,892 \cdot 10^{-1}$
313	$2,673 \cdot 10^{-1}$
323	$2,482 \cdot 10^{-1}$
333	$2,285 \cdot 10^{-1}$

En la Figura 17 se muestra la representación de los resultados de las tablas 3 y 4 para poder realizar una comparación entre los resultados calculados en este trabajo con los valores experimentales.

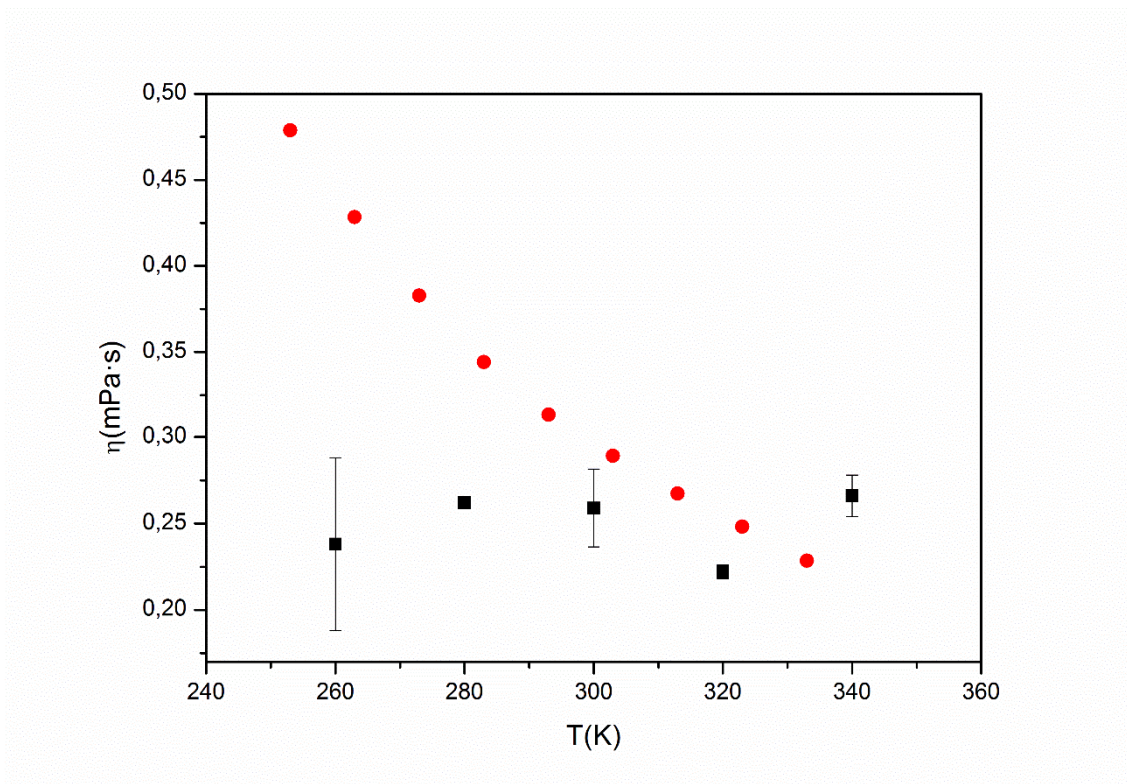


Figura 17. Representación de la viscosidad a diferentes temperaturas. Los puntos rojos representan los datos experimentales²²⁻²³. Los cuadrados negros son los valores obtenidos mediante cálculos computacionales.

Así, el comportamiento experimental observado indica que la viscosidad del hexano disminuye exponencialmente al aumentar la temperatura, algo que es de esperar². Sin embargo los resultados calculados no describen esta tendencia ya que todos ellos están alrededor de un valor de 0,25 mPa·s. Los datos a 300 y 320 K sí que reproducen de manera adecuada los valores experimentales, si bien todos los valores teóricos obtenidos están en el orden de los valores experimentales. Las diferencias observadas pueden deberse al modelo utilizado y, fundamentalmente a sus simplificaciones, como por ejemplo, la consideración del hexano como una molécula esférica.

CONCLUSIONES

- Se han aplicado con éxito las metodologías de la Mecánica y Dinámica Molecular para el estudio del coeficiente de difusión y viscosidad del hexano líquido.
- Se ha comprobado que el campo de fuerzas CHARMM36 es adecuado para el estudio de propiedades del hexano líquido a 300 K.
- Se ha observado que existe una dependencia del coeficiente de difusión calculado y el tamaño de la caja del sistema a simular, observándose que tamaños de caja pequeños subestiman el valor del coeficiente de difusión calculado.
- El aumento del número y tamaño de sistemas a simular mejoraría el cálculo del coeficiente de difusión corregido.
- Se ha encontrado que el cálculo del coeficiente de difusión varía con la asignación de las velocidades iniciales por lo que una posible mejora de la comparación de los resultados experimentales se podría obtener aumentando el número de réplicas a realizar.
- Se ha calculado la viscosidad a partir de los resultados de D aplicando el modelo de Stokes-Einstein y comparado con datos experimentales.
- La comparación de los valores de viscosidad calculados con los experimentos no es tan buena como en el caso de D , seguramente debido a que el modelo aplicado es muy sencillo.
- La consideración del hexano como esfera es una de las aproximaciones en las que este modelo se desvía del comportamiento esperado.
- Se ha logrado la ejecución con soltura de los distintos programas de cálculo y análisis que comprende el uso de GROMACS en un sistema operativo Linux y a través de líneas de comando.

ANEXOS

ANEXO 1: Archivos utilizados en el cálculo

em.mdp

```
; minim.mdp - used as input into grompp to generate em.tpr  
; Parameters describing what to do, when to stop and what to save  
Integrator = steep; Algorithm (steep = steepest descent minimization)  
emtol = 1000.0; Stop minimization when the maximum force < 1000.0 kJ/mol/nm  
emstep = 0.01; Energy step size  
nsteps = 50000; Maximum number of (minimization) steps to perform  
; Parameters describing how to find the neighbors of each atom and how to calculate the interactions  
cutoff-scheme = Verlet  
nstlist = 10 ; Frequency to update the neighbor list and long range forces  
ns_type = grid; Method to determine neighbor list (simple, grid)  
rlist = 1.0; Cut-off for making neighbor list (short range forces)  
coulombtype = PME; Treatment of long range electrostatic interactions  
rcoulomb = 1.0; Short-range electrostatic cut-off
```

rvdw = 1.0; Short-range Van der Waals cut-off

pbw = xyz; Periodic Boundary Conditions (yes/no)

npt.mdp

; Run parameters

integrator = md; leap-frog integrator

nsteps = 200000; 1 * 20000 = 400 ps

dt = 0.002; 2 fs

; Output control

nstxout = 500; save coordinates every 0.2 ps

nstvout = 500; save velocities every 0.2 ps

nstenergy = 500; save energies every 0.2 ps

nstlog = 500; update log file every 0.2 ps

; Bond parameters

continuation = yes; Restarting after NVT

constraint_algorithm = lincs; holonomic constraints

constraints = h-bonds; h-bonds constrained

lincs_iter = 1; accuracy of LINCS

lincs_order = 4; also related to accuracy

; Neighborsearching

```

cutoff-scheme = Verlet
ns_type = grid; search neighboring grid cells
nstlist = 10; 10 fs
rlist = 1.5; short-range neighborlist cutoff (in nm)
rcoulomb = 1.3; short-range electrostatic cutoff (in nm)
;rvdw = 1.5; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype = PME; Particle Mesh Ewald for long-range electrostatics
pme_order = 4; cubic interpolation
fourierspacing = 0.12
; Method for doing Van der Waals
;vdw-type = Cut-off
vdw-modifier = Potential-switch
; cut-off lengths
rvdw-switch = 1.0
rvdw = 1.3
; Temperature coupling is on

```

```

tcoupl = V-rescale; modified Berendsen thermostat
tc-grps = System; two coupling groups - more accurate
tau_t = 0.1; time constant, in ps
ref_t = 298.15; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl = Berendsen; Pressure coupling on in NPT
;pcoupl = Parrinello-Rahman; Pressure coupling on in NPT
pcoupltype = isotropic; uniform scaling of box vectors
tau_p = 2.0; time constant, in ps
ref_p = 1.01; reference pressure, in bar
compressibility = 4.5e-5; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions
pbc = xyz; 3-D PBC
; Dispersion correction
DispCorr = EnerPres; account for cut-off vdW scheme
; Velocity generation

```

gen_vel = no; Velocity generation is off

prod.mdp

; Run parameters

integrator = md; leap-frog integrator

nsteps = 250000; 2 * 50000 = 100 ps

dt = 0.001; 2 fs

; Output control

nstxout = 100; save coordinates every 0.2 ps

nstvout = 100; save velocities every 0.2 ps

nstenergy = 100; save energies every 0.2 ps

nstlog = 500; update log file every 0.2 ps

; Bond parameters

continuation = no; first dynamics run

constraint_algorithm = lincs; holonomic constraints

constraints = h-bonds; heavy atoms-H bonds constrained

lincs_iter = 2; accuracy of LINCS

lincs_order = 4; also related to accuracy

; Neighborsearching

```

cutoff-scheme = Verlet
ns_type = grid; search neighboring grid cells
nstlist = 20; 20 fs
rlist = 1.0; short-range neighborlist cutoff (in nm)
rcoulomb = 1.0; short-range electrostatic cutoff (in nm)
rvdw = 1.0; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype = PME; Particle Mesh Ewald for long-range electrostatics
pme_order = 4; cubic interpolation
fourierspacing = 0.16; grid spacing for FFT
; Temperature coupling is on
tcoupl = v-rescale; modified Berendsen thermostat
tc-grps = System; one coupling groups
tau_t = 0.1; time constant, in ps
ref_t = 260; reference temperature, one for each group, in K
; Pressure coupling is off
pcoupl = no; no pressure coupling in NVT

```



```
; Periodic boundary conditions  
pbc = xyz; 3-D PBC  
; Dispersion correction  
DispCorr = EnerPres; account for cut-off vdW scheme  
; Velocity generation  
gen_vel = yes  
gen_temp = 260; temperature for Maxwell distribution  
gen_seed = -1; generate a random seed
```

ANEXO 2: Valores del coeficiente de difusión según el tamaño del sistema y la temperatura

A continuación, se muestran los valores del coeficiente de difusión a las temperaturas de simulación (260, 280, 300, 320 y 340K).

Tabla A.1.1: Coeficiente de difusión y error a las temperaturas de trabajo calculados para la caja pequeña ($l=3,00$ nm).

Número de moléculas (N)	T (K)	D (cm^2/s)
128	260	$2,22 \cdot 10^{-5} \pm 1,66 \cdot 10^{-6}$
128	280	$2,24 \cdot 10^{-5} \pm 1,42 \cdot 10^{-6}$
128	300	$2,51 \cdot 10^{-5} \pm 1,86 \cdot 10^{-6}$
128	320	$2,03 \cdot 10^{-5} \pm 2,4 \cdot 10^{-6}$
128	340	$3,04 \cdot 10^{-5} \pm 1,56 \cdot 10^{-6}$

Tabla A.1.2: Coeficiente de difusión y error a las temperaturas de trabajo calculados para la caja mediana ($l=4,17$ nm).

Número de moléculas (N)	T (K)	D (cm^2/s)
512	260	$2,29 \cdot 10^{-5} \pm 2,36 \cdot 10^{-6}$
512	280	$2,57 \cdot 10^{-5} \pm 8,85 \cdot 10^{-7}$
512	300	$2,72 \cdot 10^{-5} \pm 2,06 \cdot 10^{-6}$
512	320	$2,9 \cdot 10^{-5} \pm 1,33 \cdot 10^{-6}$
512	340	$3,24 \cdot 10^{-5} \pm 2,01 \cdot 10^{-6}$

Tabla A.1.3: Coeficiente de difusión y error a las temperaturas de trabajo calculados para la caja grande ($l=6,00$ nm).

Número de moléculas(N)	T	D (cm^2/s)
1024	260	$2,34 \cdot 10^{-5} \pm 1,2 \cdot 10^{-6}$
1024	280	$2,72 \cdot 10^{-5} \pm 5,26 \cdot 10^{-7}$
1024	300	$3,05 \cdot 10^{-5} \pm 9,88 \cdot 10^{-7}$
1024	320	$3,14 \cdot 10^{-5} \pm 4,81 \cdot 10^{-7}$
1024	340	$3,46 \cdot 10^{-5} \pm 2,45 \cdot 10^{-6}$

ANEXO 3: Datos de ajuste para el coeficiente de difusión corregido

Se presentan las representaciones del coeficiente de difusión frente a $1/L$ a las temperaturas de trabajo estudiadas. La ordenada en el origen de estas representaciones se corresponde con el coeficiente de difusión (D). La pendiente de estas representaciones se corresponde con el error de la desviación. Por ejemplo para 260K; $D = (3,2 \pm 0,7) \cdot 10^{-5}$ y para 280K $D = (3,17 \pm 0,04) 10^{-5}$ en unidades cm^2/s .

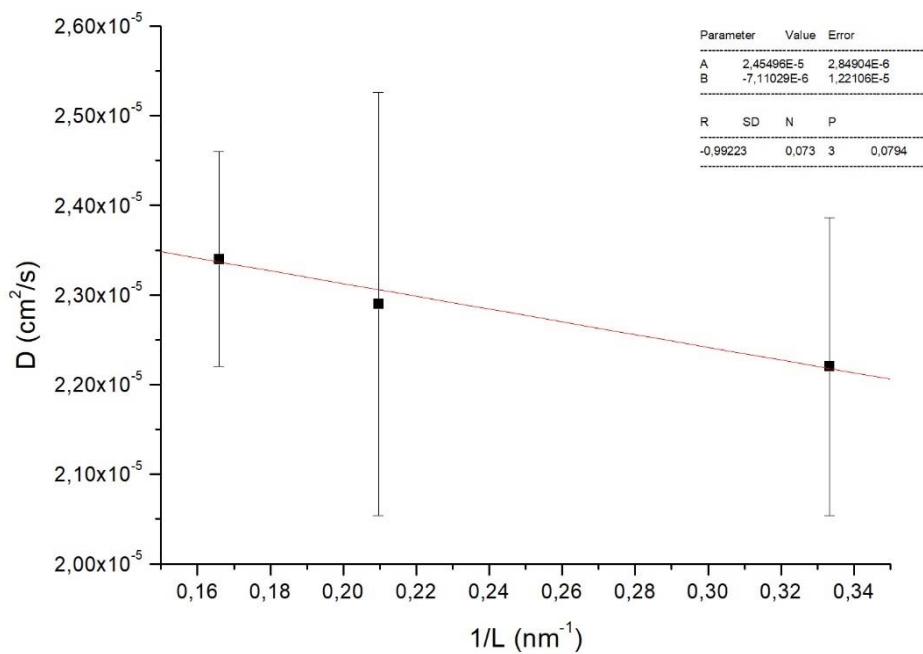


Figura A.2.1: Coeficiente de difusión frente a $1/L$ a 260K.

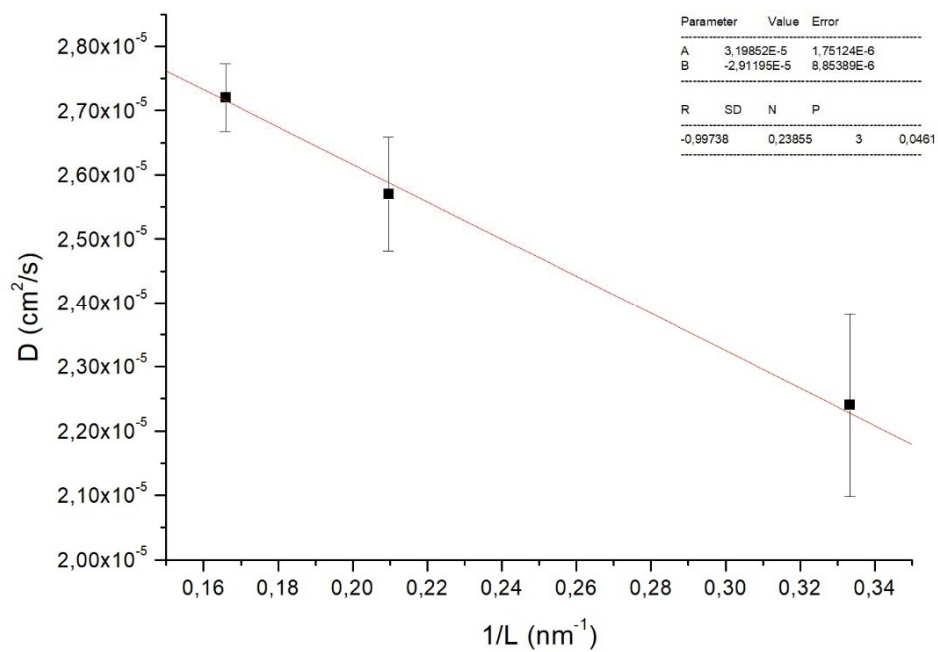


Figura A.2.2: Coeficiente de difusión frente a $1/L$ a 280K.

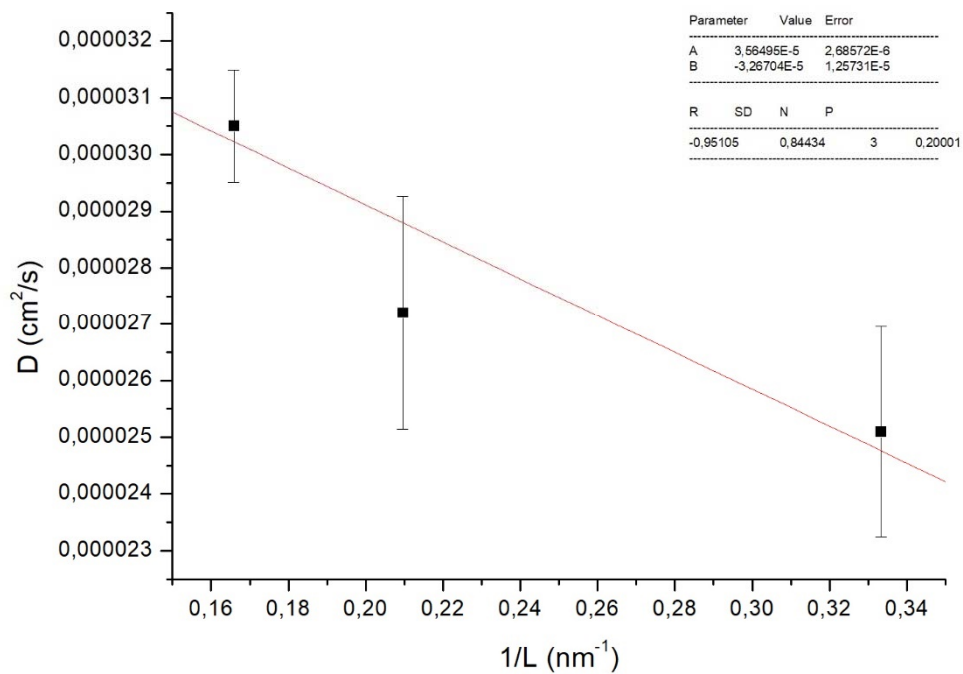


Figura A.2.3: Coeficiente de difusión frente a $1/L$ a 300K.

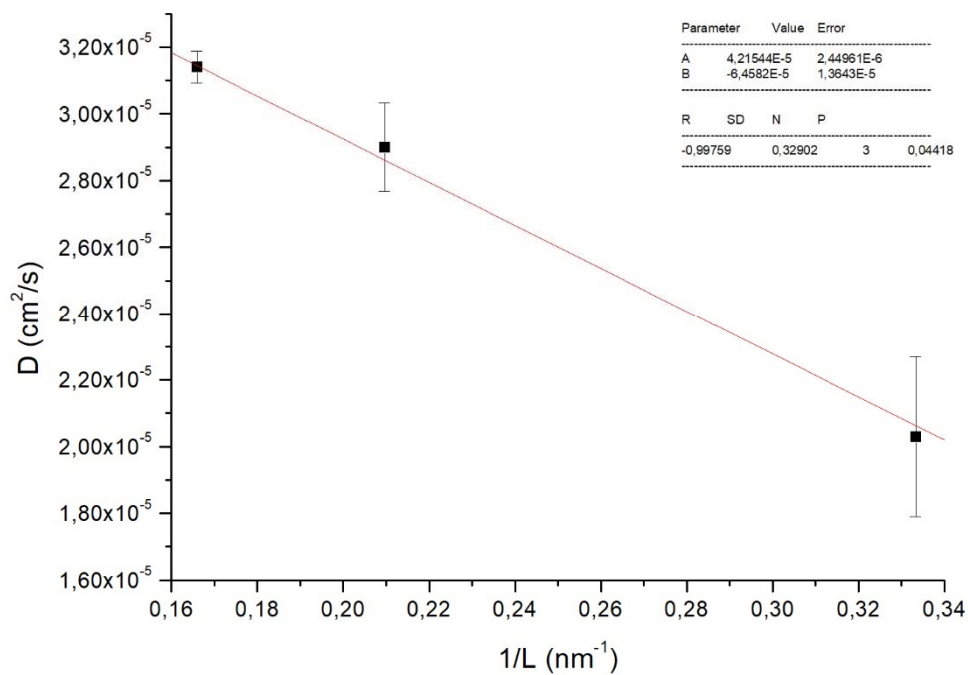


Figura A.2.4: Coeficiente de difusión frente a $1/L$ a 320K.

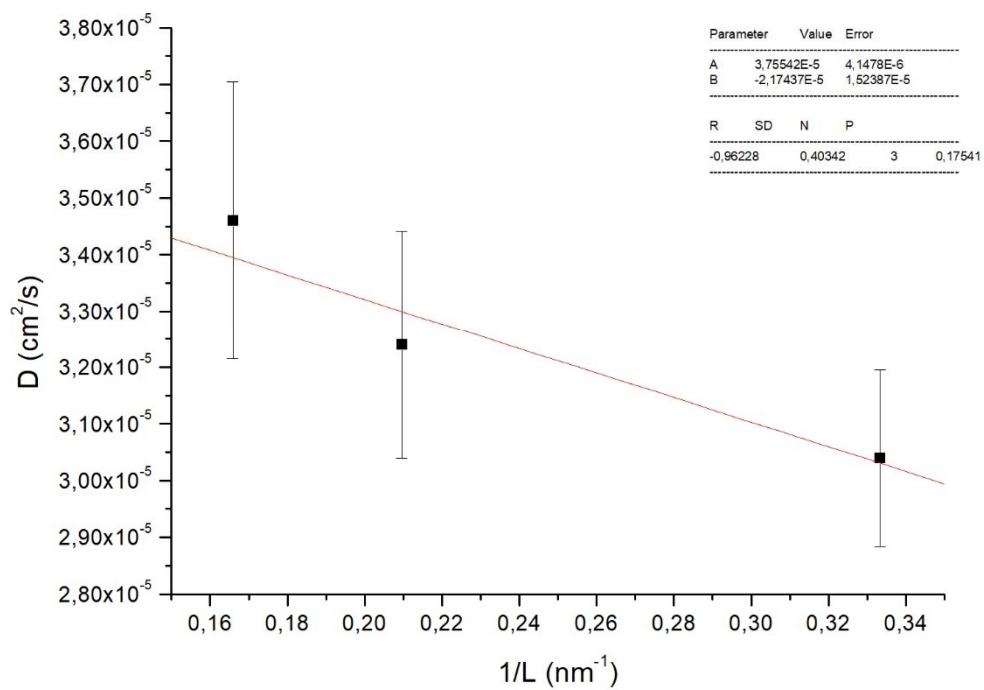


Figura A.2.5: Coeficiente de difusión frente a $1/L$ a 340K.

